

Figure 1

The novel gene as identified through RACE analysis (894 bp)

GGGAGTGGAGTGAGGGGTAACAAGATGGCGACCGAGACGGTGGAGCTCCATAAGCTAA
AGCTTGCCGAACATAAGCAAGAATGCTTGCTCGTGGTTTGGAGACCAAGGGAATAAAG
CAAGATCTTATCCACAGACTCCAGGCATATCTTGAAGAACATGCTGAAGAGGAGGCAAAT
GAAGAAGATGTACTGGGAGATGAAACAGAGGAAGAAGAAACAAAGCCCATTGAGCTCCC
TGTCAAAGAGGAAGAACCCCTGAAAAAAGTGTGATGTGGCAGCAGAGAAGAAAGTGG
TGAAAATTACATCTGAAATACCACAGACTGAGAGAATGCAGAAGAGGGCTGAACGATTCA
ATGTACCTGTGAGCTTGGAGAGTAAGAAAGCTGCTCGGGCAGCTAGGTTTGGGATTCT
TCAGTTCCAACAAAAGGCTCTGTCATCTGATAACAAACCTATGGTTAACTTGGATAAGCTG
AAGGAAAGAGCTCAAAGATTTGGTTTGAATGTCTCTTCAATCTCCAGAAAGTCTGAAGAT
GATGAGAAACTGAAAAAGAGGAAGGAGCGATTTGGGATTGTCAAGTTCAGCTGGAAC
TGGAACACAGAGGATACAGAGGCAAGAGAGGAAAAGAGCAGAGCGCTTTGGGATT
GCCTGATGAAAAGTTCCTGATACTTTCTGTTCTCCAGTGTTCCTCATTTCTCCTTCTTC
TTGGTCACATATATGCCTAAATGCACAGTCATGTGCCTACGTCCTCGCTCGCAATGAGG
GAGCATGTACCCAGGTACATCCATGAAGTGCAGCAGAGTTGACTTATTGCTGTTTCA
GCTTTAAGGTTGTTGTTTTGTTTTGATTATGTTGCTTGTTAATAAAAAAAATAGAAA

A

03783676-02240

Figure 2

Amino acid sequence as translated from the novel gene (210 amino acids)

MATETVELHKLKLAELKQECLARGLETGKIQDLIHRLQAYLEEHAEEDVVGDETEEE
ETKPIELPVKEEPEKTVDAEKKVVKITSEIPQTERMQKRAERFNVPSLESKKAARAAR
FGISSVPTKGLSSDNKPMVNLDKLKERAFGLNVSSISRKSEDEKLKKRKERFGIVTSSAG
TGTTEDEAKKRKRAERFGIA

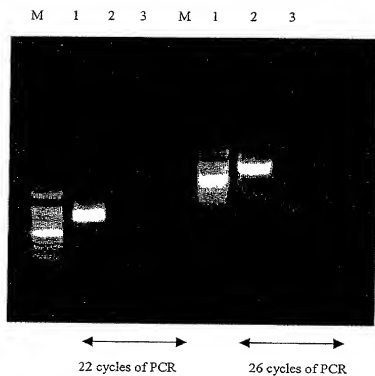
Underlined sequences are amino acid sequences obtained by MS/MS analysis.

Figure 3

The sequence of the novel gene amplified through long distant PCR and used to construct the expression vector (873 bp).

TGGAGTGAGGGGTAACAAGATGGCGACCGAGACGGTGGAGCTCCATAAGCTAAAGCTT
GCCGAACATAAGCAAGAATGTCTTGCTCGTGTTGGAGACCAAGGGAATAAAGCAAGA
TCTTATCCACAGACTCCAGGCATATCTTGAAGAACATGCTGAAGAGGAGGCAAATGAAG
AAGATGTACTGGGAGATGAAACAGAGGAAGAAGAAACAAAGCCCATTGAGCTCCCTGTC
AAAGAGGAAGAACCCCTGAAAAAACTGTTGATGTGGCAGCAGAGAAGAAAGTGGTGAA
AATTACATCTGAAATACCACAGACTGAGAGAATGCAGAAGAGGGCTGAACGATTCAATGT
ACCTGTGAGCTTGGAGAGTAAGAAAGCTGCTCGGGCAGCTAGGTTTGGGATTCTTCAG
TTCCAACAAAAGGTCTGTCATCTGATAACAAACCTATGGTTAACTTGGATAAGCTGAAGG
AAAGAGCTCAAAGATTTGGTTTGAATGTCTCTTCAATCTCCAGAAAGTCTGAAGATGATG
AGAAACTGAAAAGAGGAAGGAGCGATTTGGGATTGTCAAGTTTCAGCTGGAACGGA
ACCACAGAGGATACAGAGGCAAGAGAGGAAAAGAGCAGAGCGCTTTGGGATTGCCT
GATGAAAAGTTCCTGATACTTTCTGTTCTCCAGTGTCTTCCATTTCTCTCCTTCTTCTTGG
TCACATATATGCCTAAATGCACAGTCATGTGCCTACGTCCTGCCTCGCAATGAGGGAGC
ATGTACCCACAGGTACATCCATGAACTGCGGCAGCAGTTTGACTTATTGCTGTTTCAGCTT
TAAGGTTGTTGTGTTTTTGTGTTTGAATTATGTTGCTTGTTAAAT

Figure 4



09768126-002101

Figure 5

P-151 5'-Untranslated Region

1 75
CAGGGGCAGCAGTGAATTATCTGAACTCGGATCTTTAAAATGTGGTAGCTCTAAAGCTGATGATGTCTGGTTAGG

76 150
AAGTGGCTCTTGCCCGCCCCAGCCCCACCGCAGTTCCTTAAGCCCCGCCCATGCCCTCCAGCTTCCTCCTCA

151 225
TGTTCATCGGTTTTTTCAGGGCTCCCTTCAACGCTCCCTCTCAGTATTTAGGTCACCACTCCCTCGGCGCCCTT

226 300
TTCGCCTCCCACTTTTTCCTCAGCAACCTTACAGTCTTTGCAGCTCCTACCTGCCAGCTCAGATCCCGCTCC

301 375
GGCTATGGGCGCGGCGCGGCTACCAACCTGAAGTCTCCAGGAAGTAA CGCCTCTCCTTCTGCCCTTTCCTGT
376 450
TGGAGGAACAGAATCAGCGCTGCCACCACCCATTGGTTGGTGGTCTGTAAATGCAGAAGCACAGTTGGTTGCCATT
451 525
TCTGTCTGTTGCAAGATACAGTGCCCGCCCTCTCCAGTTCCACCTTTTGAAAGAGGTGGGGCAAGCTGCCTAG
526 600
AGAAGTGAAGAGCGAGCTCAGCTATTGACCAATGGGAAGAGCTGATGGTATGGCGTGGGAGCAAGAGTGACAACGA
601 675
TTGGTCAGCCTTGCACTCTACGCCTAAGCGGGAACTCCTGGAGGCGGAGCCGCGGGTGGGGGGAGTGGAGTG
676
AGGGGTAAACAAGATG.....P151 coding region.....

(Total length: 690 bp)

Sequence with asterisk: the 274 bp fragment

Underlined sequences are the minicistrons or uORFs before the start of the P151 coding region with the start and stop codons in bold.